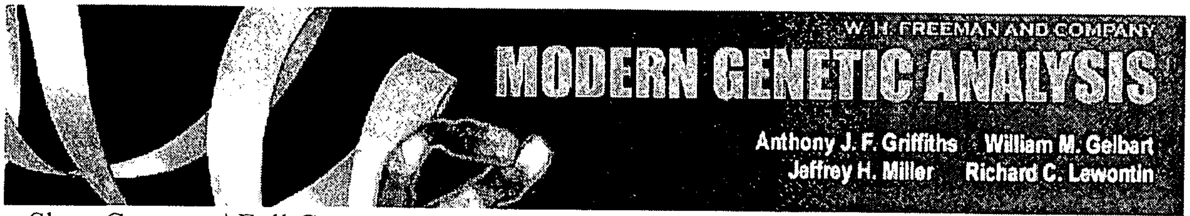


Short Contents | Full Contents

Other books @ NCBI

Modern Genetic Analysis ➔ 12. Genomics

Functional Genomics

Large-scale sequence data is the beginning of functional genomics. The following sections describe some of the analyses that can be performed to investigate function.

Characterization of the Proteome by Open Reading Frame Analysis

In ORF analysis, the genomic DNA sequence is fed into a computer, which examines each of the six reading frames of all chromosomes and searches for segments beginning with the translational start codon AUG and ending with a stop codon. Any ORFs of at least 100 codons are candidates for genes. Most ORFs will be completely novel, not corresponding to any familiar gene with alleles producing identifiable phenotypes. The ORFs can be analyzed for function initially by using the computer to search data bases looking for full or partial homology to known genes characterized in other organisms. The location, orientation, and clustering of ORFs also are important genomic information. Examples from *Hemophilus* and *Saccharomyces* are shown in Figures 12-33 on page 400 and 12-34 on page 401. A provisional proteome gene distribution can be deduced from such analysis (see examples in Chapter 3). 🡅 TOP

Gene Disruption: Knockouts

ORF function can be investigated by systematically knocking out the gene by in vitro mutagenesis and then looking for any possible mutant phenotype that might provide clues about function. This process is underway in the fully sequenced genomes. Interestingly, many ORFs, when knocked out, show no phenotypic effects; such ORFs are called "orphan" genes. Others are shown to be essential because, when knocked out, lethality ensues. 🡅 TOP

The Study of Gene Interactions by Suppressor Analysis

A mutant allele of one gene can often be suppressed or modified by a range of mutations in other genes (see Chapter 6). Elucidation of such patterns of suppression and epistasis generates networks of interacting proteins. 🡅 TOP

The Study of Gene Interactions by the Yeast Two-Hybrid System

This method investigates interaction by using a twoplasmid system in yeast. The basis for the test is the yeast GAL4 transcriptional activator. This protein has two domains, a DNA-binding domain and an activation domain, both of which must be

in close juxtaposition in order for the protein to initiate transcription. A gene for one protein under investigation is spliced next to the GAL4 DNA-binding domain on one plasmid, and this protein acts as "bait." On another plasmid, a gene for another protein under test is spliced to the activation domain; this protein is said to be the "target" (Figure 12-35 on the next page). The two plasmids are then introduced into the same cell. One way of doing this is to mate haploid cells containing bait and target. The only way that the GAL4 binding and activation domains can come together is if the bait and target proteins bind to each other, demonstrating a physical interaction. The two-hybrid system can be automated to facilitate large-scale hunting for protein interactions throughout the proteome. ↑ TOP

The Study of Developmental Regulation by Using DNA Chips

DNA chips are about to revolutionize genetics in the same way that silicone chips revolutionized the computer industry. DNA chips are samples of DNA laid out in regimented arrays bound to a glass "chip" the size of a microscope cover slip.

One protocol is as follows. Robotic machines with multiple printing tips resembling miniature fountain pen nibs deliver microscopic droplets of DNA solution to specific positions (addresses) on the chip. The DNA is dried and treated so that it will bind to the glass. Thousands of samples can be applied to one chip. In one protocol, the array of DNAs consists of known cDNAs from different genes. In principle, all the cDNAs of an entire genome could be arrayed on a small number of chips. The chips are exposed to a heterogeneous, developmentally specific probe. The probe can be either labeled total mRNA extracted at some specific stage of development or cDNA made from the total mRNA. Fluorescent labels are used, and the binding of the probe molecules to the glass chip is monitored automatically with the use of laser beams. A typical result is shown in Figure 12-36a. In this way, the genes that are active at any stage of development or under any environmental condition can be assayed. Once again, the idea is to identify protein networks that are active in the cell at any particular stage of interest. Figure 12-37 shows an example of a developmental expression sequence.

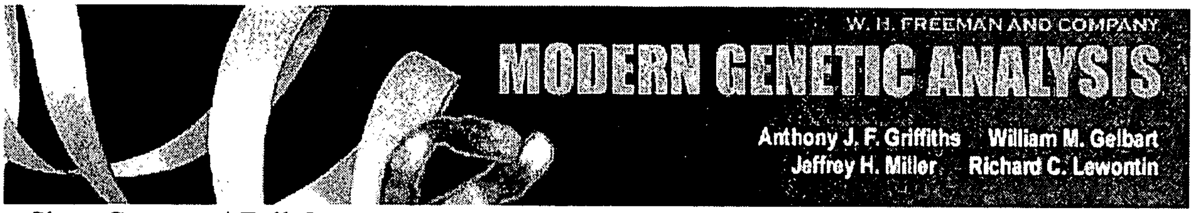
Another protocol loads the chip with an array of oligonucleotides actually synthesized nucleotide by nucleotide on the chip itself (Figure 12-38 on page 404). The glass is first covered with protecting groups that prevent DNA deposition. A mask is placed on the glass with holes corresponding to the sites of deposition. Then laser beams are shone onto the holes where synthesis is to begin. The light knocks off the protecting groups. Then the glass is bathed in the first nucleotide to be deposited. Each nucleotide carries its own protection group, which can be knocked off for the second round of deposition. Hence, by sequential application of the appropriate masks and bathing sequences, arrays of different nucleotides can be built up. For studying genomic function, these oligonucleotides could be identifying sequences of genes, such as ESTs, or characteristic sequences. As before, the completed array is bathed in fluorescent probe isolated at some developmental stage. Binding to an oligonucleotide array is shown in Figure 12-36b.

The oligonucleotide procedure was used to produce a complete array of the genome of baker's yeast. For each of the 6200 genes in the genome, 20 oligomers of 25 bases were made, on the basis of the known genomic sequence. These oligonucleotides were arrayed on four chips in the same order as the sequences are on the

chromosomes. The array was used to determine differences in gene expression between cells grown in various physiological conditions. Probes were made from total mRNA isolated from cells grown under the different conditions.

Note that these DNA array methods basically take an approach to genetic dissection that is an alternative to mutational analysis. Under either method the goal is to define the set of genes or proteins that are important to any specific process under study. Traditional mutational analysis accomplishes this objective by amassing mutations that affect a specific process under study; chip technology does it by detecting the specific mRNAs that are transcribed in that process.

DNA chips can also be used to detect mutations. Oligonucleotides can be prepared that are complementary to all possible simple mutational changes in a genetic region under analysis. Alternatively, oligonucleotides complementary to all the known mutations in a human gene (such as a breast cancer gene) can be arrayed on the chip. ↑ TOP



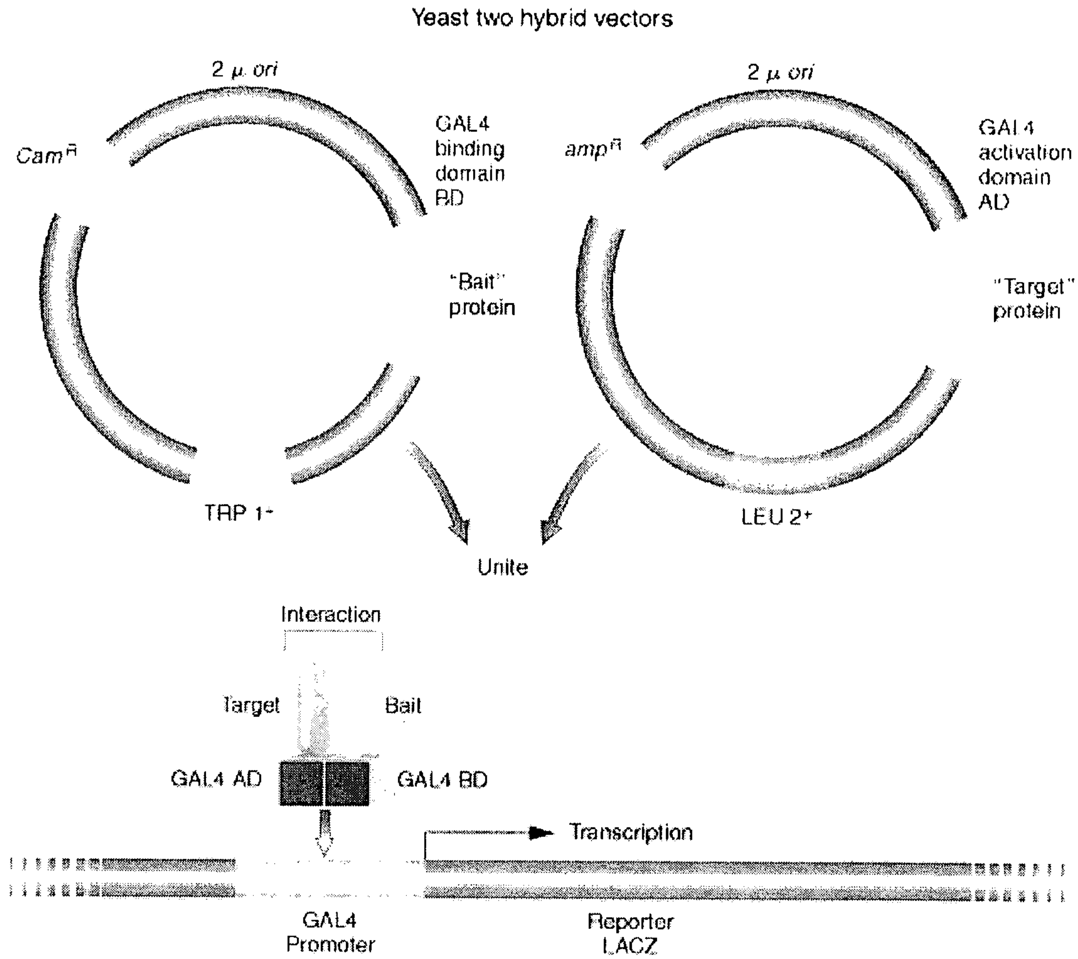


Figure 12-35. The yeast two-hybrid system for detecting gene interaction. The system uses the binding of two proteins under test to restore the function of the GAL4 protein, which activates a reporter gene.